УДК 519.71;612.82

NEURAL NETWORKS WITH EVOLVING SYNAPTIC CONNECTIONS

A.D. LINKEVICH
(Polotsk State University)

Analogous neural networks are considered for which interneuronal (synaptic) connections evolve in time according to some dynamical law. Such a dynamics have been found for the synaptic efficacies that the corresponding network performs nonlinear optimization. Minimizing function can be chosen in such a way that learning and retrieval of patterns are easily done. Neural networks are designed able to self-learning, i.e. parameters of such a system are adjusted due to an intrinsic network dynamics so as to ensure learning signals received and no external agents (teachers) are necessary to train the network. The above networks can be supplemented by a specific neural network with evolving synaptic couplings that function as a detector of novelty in such a way that an incoming signal is kept intact, if it is linearly independent on patterns received by the system in the past, and put it to zero otherwise.

A recent advance in grasping mechanisms of information processing by neural systems is considerably affected by the known paper of Hopfield [1] paved the way for analytical treatment such systems due to concepts and methods of statistical physics and other theoretical tools since Hamiltonian-like functions have been introduced both for binary [1] and analogous [2, 3] neural networks. Most of works in this research area treat neural networks as systems characterized by the only type of dynamical variables – states of neurons, while synaptic connections between neurons are taken to be rigid (constant in time) during the network evolution. (The synaptic efficacies are modified in a learning process but the latter is usually considered separately from the network dynamics).

Although a certain progress has been achieved due to the Hopfield model (see, e.g., [4]), a huge number of problems is still open on the way towards both understanding brain mechanisms and designing neuro-computers. One possible way out is to study generalizations of the Hopfield models. Some results along such a line of modifications as analogous networks with asymmetric interneuronal connections are given in [5]. The direction of improving facilities of neural networks adopted in this paper is exploring analogous networks with symmetric synaptic efficacies which evolve in time in accordance with some dynamical law.

Some works have been done in which networks have flexible synapting couplings. So, investigations of the so-called unsupervised learning are aimed to describe how synapses evolve in time with available local signal information (see, e.g., [6] and references therein). In the area of pattern recognition, e.g., such systems adaptively cluster patterns into classes. Learning and recognition of temporal sequences of patterns were studied in a model with randomly delayed interactions [7]. A «brainwashing» algorithm have been proposed [8] for modification of interneuronal connections in a context of development and self-organization of neural networks.

The approach of the present work consists of three stages. First I find such a dynamics of synaptic connections that the corresponding network fulfils nonlinear optimization, i.e. minimization of a given arbitrary nonlinear function. This task is interesting itself and important in view of a broad area of applications. (For another approach to nonlinear optimization by electrical circuits which are close to neural networks see [9] and references therein). The second step is application of the above networks to associative memory goal. Namely, I show that functions for minimization can be chosen in such a way that learning and retrieval of patterns are easily done. Further I design systems able to adaptive learning or self-learning. This means that parameters of such a network are adjusted due to the usual (intrinsic) network evolution so as to ensure learning signals received by the system and no external agents are necessary to train the network. Finally a specific network is proposed with evolving connections that functions as some kind of detector of novelty selecting incoming signals which are linearly independent on signals received in the past.

Consider a network composed of N neuron-like elements whose states evolve in time according to the equations [3]:

$$\dot{u}_i = -\gamma_i u_i + \sum_{j=1}^{N} T_{ij} \upsilon_j + I_i , \qquad i = \overline{1, N} . \tag{1}$$

Here the variable $u_i = u_i\ t$ describes the state of neuron i at time t, the positive parameter γ_i characterizes its properties, the synaptic efficacy T_{ij} determines strength of influence of the output of the j-th neuron υ_j on the input of neuron i through their interneuronal (synaptic) connection, I_i is some any other (e.g., sensory) input to neuron i. The neuronal variables υ_i and u_i are connected by an instantaneous input-output transfer function $f_i(\cdot)$, i.e. $\upsilon_i(t) = f_i[\lambda u_i(t)]$.

This function is normally taken to be a monotone sigmoid and therefore positive parameter λ controls its maximal slope. As is shown in [3] (see also [2]) the quantity

$$\varepsilon = -\frac{1}{2}\sum_{ij} T_{ij}\upsilon_i\upsilon_j - \sum_i I_i\upsilon_i + \frac{1}{\lambda}\sum_i \gamma_i \int_0^{\upsilon_i} d\upsilon f_i^{-1}(\upsilon) \tag{2}$$

is a Hamiltonian-like function for the system, i.e. it plays the role of a global Lyapunov function, if the matrix T_{ij} is symmetric and constant in time.

In contrast to [2, 3], I consider the case when the T_{ij} and I_i evolve in time in such a way that

$$\dot{T}_{ij} = \alpha G_i G_j, \quad \dot{I}_i = \mu G_i, \quad i, j = \overline{1, N}. \tag{3}$$

Here $G_i = \partial\varphi(\vec{\upsilon})/\partial\upsilon_i$ is a component of the gradient of some (nonlinear) function $\varphi(\vec{\upsilon})$, $\vec{\upsilon} = (\upsilon_1,\ldots,\upsilon_N)$; λ is some positive constant or function and $\mu = \mu(\xi)$, is such a function of the scalar product $\xi = \vec{G}\vec{\upsilon}$ that the conditions $\mu(\xi)\cdot\xi \ge 0$, $\forall\xi \in R$ and $\mu(0) \ne 0$ are satisfied. Then it is easy to prove that the function ε is again a Lyapunov function for the system and $\dot{\varepsilon} = 0$ if $\vec{\upsilon} = const$ and $\vec{G} = \nabla\varphi(\vec{\upsilon}) = 0$. This means that our network converges to such a stationary point that the necessary condition for a local extremum of the function $\varphi(\vec{\upsilon})$ is held. Therefore such a network can be used for nonlinear optimization.

The discrete-time counterpart to the system (1) writen in the form [10]

$$\frac{u_i^{t+1}}{R_i} = \sum_{j=1}^{N} T_{ij}^t \upsilon_j^t + I_i^t, \qquad i = \overline{1, N},.. \quad t = 0,1\ldots \tag{4}$$

can be treated as well. Indeed, introducing new variables x_i^t by the relations $\sum T_{ij}^{t-1} x_j^t = u_i^t / R_i - I_i^{t-1}$ and redefining the T_{ij} and I_i so as $T_{ij}^{t-1} \to T_{ij}^t$ and $I_{ij}^{t-1} \to I_i^t$, we recast eqs.(4):

$$x_i^t = F_i\left(\sum_j T_{ij}^{t-1} x_j^{t-1} + I_i^{t-1}\right), \qquad i = \overline{1, N}, \qquad t = 0,1\ldots \tag{5}$$

where $F_i(z) = f_i(R_i z)$ with z dummy argument. Take the discrete-time analogs of eqs. (3)

$$T_{ij}^t = T_{ij}^{t-1} + \alpha^t G_i^t G_j^t, \qquad I_i^t = I_i^{t-1} + \mu^t G_i^t, \tag{6}$$

where $G_i^t = \partial\varphi(\vec{x})/\partial x_i\big|_{\vec{x}=\vec{x}^t}$. Then we can construct the global Lyapunov function for the system:

$$\varepsilon^t = -\frac{1}{2}\sum_{ij} T_{ij}^t x_i^t x_j^t - \sum_i I_i^t x_i^t + \int_0^{x_i^t} dz F_i^{-1}(z),$$

which is similar to eq.(2). We find that $\Delta\varepsilon^t \le 0$ and $\Delta\varepsilon^t = 0$ if $\vec{x}^t = const$ and $\vec{G}^t = \nabla\varphi(\vec{x})\big|_{\vec{x}=\vec{x}^t}$ for all time moments t after some t^*. (For more detail about nonlinear optimization by the networks above formulated, some possible applications and ways of hardware implementation of the networks see [11]).

Now turn to the problem of autoassociative memory. Suppose that a prescribed set of patterns represented by the vectors of the network state $\vec{\xi}^1,\ldots,\vec{\xi}^p$ should be memorized. Then it is enough to construct such a function $\varphi(\vec{x})$ that every $\vec{\xi}^\mu$ is a local minimum of $\varphi(\vec{x})$.

It is obvious that this task can easily be solved in a number of ways. E.g., we can take such quadratic functions as

$$\varphi(\vec{x}) = \sum_{\mu} c^{\mu}\left(\vec{x}-\vec{\xi}^{\mu}\right)^{2}, \text{ or } \varphi(\vec{x}) = \sum_{\mu}\left(\vec{x}-\vec{\xi}^{\mu}\right)^{T} Q\left(\vec{\xi}^{\mu}\right)\left(\vec{x}-\vec{\xi}^{\mu}\right) \tag{7}$$

with appropriately chosen c^{μ} and $Q\left(\vec{\xi}^{\mu}\right)$.

Mention main features of such a memory: (i) Number of memorized patterns p can be arbitraryly large (cf. with the Hopfield model for which $p < N$ [4]). (ii) Basins of attraction around the patterns and rate of convergence towards them can be controlled (to some extent, of course). (iii) Full phase space of such a network can be used for storage of information. All these properties are achieved by appropriate «gardening» of the landscape of the minimizing function $\varphi(\vec{x})$.

Now we are in a position to design systems able to adaptive learning or self-learning. Indeed, consider a network which evolve, e.g., in discrete time in such a way that its dynamics is given by the equations

$$x_i^t = F_i\left(\sum_j T_{ij}^{t-1} x_j^{t-1} + I_i^{t-1} + \rho_i^t\right), \tag{8}$$

$$\mathit{\Gamma}_i^t(\vec{z}) = \kappa^t \mathit{\Gamma}_i^{t-1}(\vec{z}) + \gamma_i^t\left(\vec{z}, \vec{\xi}^t\right) \tag{9}$$

and by eqs. (6) in which $G_i^t = \mathit{\Gamma}_i^t\left(\vec{x}^t\right)$. Here ρ^t and ξ^t are input (sensory) signals incoming to the system (they can be related with each other; e.g., a simple reasonable variant is $\xi_i^t = F_i\left(\rho_i^t\right)$; κ^t is a constant or function restricted by the condition $0 < \kappa^t < 1$. The variable $\vec{z}$ belongs to the phase space of the network so that eq.(9) together with the assignment $G_i^t = \mathit{\Gamma}_i^t\left(\vec{x}^t\right)$ means iterative computation of the function $\vec{\mathit{\Gamma}}(\vec{z})$ which brings the current value of the gradient for the motion of the system. A simple underlain picture is that our network evolves in such a way that a minimizing function $\varphi(\vec{z})$ is iteratively calculated according to the relation

$$\varphi^t(\vec{z}) = \kappa^t \varphi^{t-1}(\vec{z}) + \psi\left(\vec{z}, \vec{\xi}^t\right). \tag{10}$$

Then

$$\mathit{\Gamma}_i^t(\vec{z}) = \partial\varphi^t(\vec{z}) \big/ \partial z_i, \qquad \gamma_i^t\left(\vec{z}, \vec{\xi}^t\right) = \partial\psi\left(\vec{z}, \vec{\xi}^t\right) \big/ \partial z_i.$$

It is obvious that the increment term $\psi\left(\vec{z}, \vec{\xi}^t\right)$ should be chosen so as to ensure a required form for the $\varphi^t(\vec{z})$ (in particular, $\psi^t(\vec{z}, 0) = 0$). E.g., in order to obtain the quadratic function (7), we put

$$\psi\left(\vec{z}, \vec{\xi}^t\right) = \left(\vec{z}-\vec{\xi}^t\right)^T Q\left(\vec{\xi}^t\right)\left(\vec{z}-\vec{\xi}^t\right)$$

and therefore

$$\gamma_i^t\left(\vec{z}, \vec{\xi}^t\right) = 2\sum_k Q_{ik}\left(\vec{\xi}^t\right)\left(z_k - \xi_k^t\right).$$

Analyse behaviour of a network described by eqs. (8), (9), (6). First consider the case when $\vec{\rho} = 0$ but $\vec{\xi} \neq 0$. We see that forming the function $\vec{\mathit{\Gamma}}(\vec{z})$ takes place, i.e. this regime is learning the system (iterations of $\vec{x}$, T and $\vec{I}$ are idle).

A second particular situation is $\vec{\xi} = 0$ (but $\vec{\rho} \neq 0$). Since $\kappa^t < 1$ and $\psi(\vec{z}, 0) = 0$, then eq. (9) yields a slow decay of the $\vec{\mathit{\Gamma}}(\vec{z})$, i.e. forgetting the stored information (this phenomenon presents permanently in our model).

Suppose that incoming signal $\vec{\rho}$ is such that the network state $\vec{x}$ becomes close to a memorized pattern $\vec{\xi}^{\mu}$ (after some time). Then, due to the intrinsic dynamics of the system discussed above, the state converges to the fixed point $\vec{\xi}^{\mu}$ This is nothing but retrieval of stored information.

In a general case when both $\vec{\rho} \neq 0$ and $\vec{\xi} \neq 0$ such a network displays a more complex behaviour accompanied by learning, forgetting, remembering and other transforming patterns.

The networks above discussed may be improved in different ways. E.g., we can supplement such a system by some kind of detector of novelty filtering input information. So, let an incoming signal $\vec{\xi}$ be transformed so as $\vec{\xi}_i^t \to \eta^t \xi_i^t$, where $\eta^t = \theta(\vec{\xi}^t \cdot \vec{d}^t)$, $\theta(x) = 1$, if $x > 0$ and $\theta(x) = 0$ if $x \leq 0$, $\vec{d}^t = D^t \vec{\xi}^t$ and matrix D evolves in time in accordance with the iterated map

$$D_{ij}^{t+1} = D_{ij}^t - \frac{d_i^t d_j^t}{\vec{\xi}^t \cdot \vec{d}^t}, \qquad D_{ij} = \delta_{ij}, \qquad i, j = \overline{1, N}.$$

Then using results of [12] one can show that the input signal is kept intact if it is linearly independent on the previous signals and the $\vec{\xi}^t$ is changed to zero if it is a linear combination of signals received by the network in the past. The system just above described can be viewed as some kind of neural networks whose synaptic efficacies D_{ij} evolve in time.

The results above reported clearly show that generalization of neural networks to systems with evolving interneuronal connections open new possibilities for information processing.

REFERENCES

1. Hopfield J.J. Proc. Natl. Acad. Sci. USA, 1982, 79, 2554.
2. Cohen M.A., Grossberg S. IEEE Trans. Syst. Man, Cybern., 1983, SMC-13, 815.
3. Hopfield J.J. Proc. Natl. Acad. Sci. USA, 1984, 81, 3088.
4. Amit, D.J. Modelling Brain Functions / D.J. Amit. – Cambridge: Cambridge Univ. Press, 1989.
5. Kartynnick, A.V. Optical Memory and Neural Networks / A.V. Kartynnick, A.D. Linkevich, 1994. – 329 p.
6. Kosko B. IEEE Trans. Neural Networks, 1990, 1, 44.
7. Bauer K., Krey U. Z.Phys.B., 1990, 79, 4761.
8. Clark J.W., Winston J.V., Rafelski J. Phys. Lett. A, 1984, 102, 207; Clark J.W. Phys. Rep., 1988, 158, 91.
9. Kennedy M.P., Chua L.O. IEEE Trans. Circuits Syst., 1988, 35, 554.
10. Marcus C.M., Westervelt R.M. Phys. Rev. A, 1989, 40, 501.
11. Linkevich, A.D. Neural networks and nonlinear optimization / A.D. Linkevich. – Novapolatsk, 1992. – (Preprint MIGMA).
12. Linkevich A.D. J. Phys. A, 1992, 25, 4139.

Поступила 15.02.2010